*Journal of Advances in Medical and Pharmaceutical Sciences*

*Volume 28, Issue 9, Page 84-103, 2026; Article no.JAMPS.163511*
*ISSN: 2394-1111*

# Osteogenesis Imperfecta from Historical Nosology to Gene-directed Therapy: A Critical Narrative Review of Mechanism, Evidence and Translational Uncertainty

**Stefan Bittmann [a++*], Elisabeth Luchter [a] and Elena Moschüring-Alieva [a#]**

*[a] Ped Mind Institute, Department of Pediatrics, Hindenburgring 4, D-48599 Gronau, Germany.*

***Authors' contributions***

*This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.*

***Article Information***

DOI: https://doi.org/10.9734/jamps/2026/v28i9890

**Open Peer Review History:**
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://pr.sdiarticle5.com/review-history/163511

***Review Article***

***Received: 03/07/2026***
***Accepted: 23/08/2026***
***Published: 26/08/2026***

**Abstract**

Osteogenesis imperfecta is the prototypical heritable disorder of bone fragility, and the last three decades have transformed it from a clinically defined syndrome into a molecularly stratified group of conditions involving more than twenty genes. That transformation has not been matched by comparable progress in classification, outcome measurement or therapeutic evidence, and the resulting mismatch is the central concern of this review. The purpose of the article is to evaluate critically how historical nosology, molecular pathology and emerging gene-directed interventions relate to one another, and to determine which conclusions the accessible evidence can currently sustain. Literature was identified through open scholarly databases, citation searching and a clinical trials registry, with appraisal directed at study design, cohort ascertainment, endpoint selection and the distance between mechanistic plausibility and demonstrated clinical benefit. Three critical findings emerge. First, competing nosological frameworks now coexist without prospective validation against clinical outcomes, so classification supports communication more than it supports prognosis or treatment selection. Second, the dominant-negative mechanism that accounts for most disease produces a hypermineralised yet mechanically inferior matrix, which explains why densitometric gains have

[++] *Professor;* [#] *Advisor;*
**Corresponding author: E-mail: stefanbittmann@gmx.de;*

***Cite as:*** *Bittman, S., Luchter, E., & Moschüring-Alieva, E. (2026). Osteogenesis Imperfecta from Historical Nosology to Gene-directed Therapy: A Critical Narrative Review of Mechanism, Evidence and Translational Uncertainty. Journal of Advances in Medical and Pharmaceutical Sciences, 28(9), 84–103. https://doi.org/10.9734/jamps/2026/v28i9890*

repeatedly failed to translate into proportionate fracture reduction and why surrogate endpoints have misdirected therapeutic expectation. Third, cell and gene therapy remain overwhelmingly preclinical or exploratory: published editing work is largely confined to cell models and rodent studies, mesenchymal stromal cell evidence rests on small uncontrolled series, and no completed clinical gene therapy trial for this condition was identified. Major unresolved questions concern the qualification of surrogate endpoints, the penetrance of pathogenic variants in unselected populations, allele-specific correction of private dominant-negative variants, and durable delivery to skeletal progenitors in a growing skeleton. Progress will depend less on new mechanistic detail than on trial designs and outcome measures capable of detecting clinically meaningful change in a rare, heterogeneous and slowly evolving disorder.

*Keywords: Osteogenesis imperfecta; type I collagen; COL1A1; gene editing; bone material quality; surrogate endpoints; rare disease trial design; sclerostin inhibition.*

## 1. Introduction

Osteogenesis imperfecta denotes a group of heritable disorders in which reduced bone mass and abnormal bone material combine to produce fragility that is disproportionate to any applied load. The clinical spectrum extends from perinatal lethality through severe progressive deformity to adults who sustain few fractures and are identified only after a family member receives a diagnosis. Bone fragility dominates clinical attention, yet the condition also affects teeth, hearing, the cardiovascular and respiratory systems, muscle and connective tissue more generally, which reflects the ubiquity of type I collagen in the extracellular matrix (Forlino & Marini, 2016; Marini *et al*., 2017). Registry data from the Netherlands place the median annual incidence at approximately 6.5 per 100 000 live births and demonstrate a hospitalisation rate roughly threefold that of the general population, rising to more than eightfold in individuals below twenty years of age (Storoni *et al*., 2022). French national hospital discharge data indicate that mortality is concentrated in the neonatal period, with immediate mortality of about 30 per cent among newborns receiving a first inpatient diagnosis, compared with under one per cent among children diagnosed at two years or older (Philippoteaux *et al*., 2025). These figures locate the condition among rare diseases with substantial and unevenly distributed burden. Recent syntheses similarly describe osteogenesis imperfecta as a genetically heterogeneous disorder with variable skeletal and extraskeletal manifestations across the clinical spectrum (Hasegawa, 2025).

The intellectual history of the disorder is unusually instructive. Descriptions of congenital bone fragility accumulated through the eighteenth and nineteenth centuries, and the association of fragility with blue sclerae, dentinogenesis imperfecta and progressive hearing loss was recognised long before any molecular explanation existed. The decisive conceptual step came with the recognition that the apparent clinical continuum concealed genetic heterogeneity, which supported a division into four clinical types distinguished by inheritance pattern, radiographic features, scleral hue and the presence of dentinogenesis imperfecta (Sillence *et al*., 1979). That framework organised the field for a generation and remains embedded in clinical vocabulary, therapeutic trials and patient advocacy.

Molecular genetics subsequently dismantled the assumption on which the framework implicitly rested, namely that clinical type would prove to be a reliable proxy for cause. Structural and quantitative defects of the α1 and α2 chains of type I collagen were shown to underlie the great majority of cases, and recessive forms attributable to collagen-modifying, chaperone, secretory and signalling proteins were subsequently identified (Cabral *et al*., 2007; Jovanovic *et al*., 2022). The gene list has continued to lengthen, and current genetic nosologies enumerate more than twenty distinct entities (Jovanovic & Marini, 2024; Unger *et al*., 2023). Clinical severity, however, distributes across genes rather than mapping onto them, and the numbered types that persist in clinical use no longer correspond to discrete molecular mechanisms. Recent molecular cohort evidence further demonstrates variation across dominant and autosomal recessive forms, reinforcing the limited predictive value of a simple clinical type-to-gene correspondence (Sait *et al*., 2025).

Therapeutic development has followed a narrower path. Cyclical intravenous bisphosphonate administration produced striking increases in bone mineral density and apparent clinical improvement in severely affected children, and antiresorptive treatment became standard practice within a decade of its introduction (Glorieux *et al*., 1998; Rauch & Glorieux, 2004). Subsequent controlled evidence has been considerably less decisive about fracture outcomes than about densitometric ones (Dwan *et al*., 2016). Anabolic and pathway-targeted approaches, including sclerostin inhibition and transforming growth factor beta blockade, were designed to address the anabolic deficit rather than to suppress resorption, and each has generated encouraging early-phase data (Glorieux *et al*., 2024; Song *et al*., 2022). Cell and gene-directed strategies aim further upstream, seeking to replace defective osteoprogenitors or to correct the causative variant, and their preclinical literature has expanded rapidly (Kocsy *et al*., 2025; Schindeler *et al*., 2022). Recent reviews also describe a widening therapeutic landscape that includes antiresorptive, anabolic, cell-based and gene-directed strategies, while the strength of clinical evidence remains uneven across these approaches (Chaugule *et al*., 2025).

Several tensions run through this literature and remain unresolved. Classification schemes have proliferated faster than evidence that any of them improves clinical decision-making. Genotype-phenotype relationships derived from referral cohorts are being challenged by population-scale sequencing that suggests substantially reduced penetrance. Densitometry continues to serve as the principal efficacy endpoint despite mechanistic reasons to distrust it in a disorder of bone material quality. Preclinical gene therapy findings are frequently described in translational language that the underlying experimental designs do not yet support. Recent synthesis of haploinsufficiency-associated disease further illustrates that descriptive severity labels may obscure clinically relevant heterogeneity within a molecularly defined subgroup (Sclocco *et al*., 2026).

### 1.1 Scope, Research Gap and Objectives

This review addresses osteogenesis imperfecta as a disorder of type I collagen biology and of the wider osteoblast network that produces, modifies and secretes collagen, and it treats historical classification, molecular pathology, outcome measurement, pharmacological management and gene-directed therapy as a single connected problem rather than as separate literatures. The population of interest comprises children and adults with a clinical or molecularly confirmed diagnosis, together with the animal, cellular and induced pluripotent stem cell models used to investigate mechanism and to test intervention.

Existing reviews are extensive but partitioned. Genetic reviews catalogue causative genes without interrogating whether the resulting nosology serves clinical purposes. Therapeutic reviews summarise agents and effect sizes without confronting the surrogate-endpoint problem that limits interpretation of almost every trial in the field. Reviews of cell and gene therapy describe technical platforms in detail while giving comparatively little attention to the delivery, durability and regulatory obstacles that separate a corrected cell line from a treated child. The specific gap addressed here is the absence of an integrated critical account that traces a single argument from the historical origins of classification, through the mechanistic basis of material-level bone failure, to the evidentiary standards that gene-directed therapies will have to meet.

The review deliberately excludes detailed orthopaedic surgical technique, rehabilitation protocols, anaesthetic management and the differential diagnosis of inflicted skeletal injury. These are substantial fields with their own methodological literatures, and superficial coverage would add length without analytical value. Discussion of individual variant interpretation is limited to what is necessary to evaluate genotype-phenotype claims.

The objectives are to evaluate the coherence and clinical utility of competing classification frameworks; to assess how far molecular mechanism explains phenotypic variability and how far it does not; to examine why densitometric improvement has repeatedly failed to predict fracture outcomes and what this implies for endpoint selection; to appraise the strength of evidence supporting current pharmacological practice; to determine the actual maturity of cell and gene therapy evidence as distinct from its rhetorical presentation; and to derive prioritised research recommendations that follow from the identified weaknesses.

## 2. Methods for Literature Selection

### 2.1 Review Design and Rationale

A critical narrative design was selected because the review question is interpretive rather than estimative. The literature spans clinical genetics, matrix biology, bone material science, paediatric endocrinology and gene therapy, and it combines randomised trials, observational cohorts, registry analyses, case series, animal experiments and cell models. No common effect measure links these designs, and the interventions of interest differ in mechanism, target population and stage of development, so a meta-analytic summary would either exclude most of the relevant evidence or aggregate incommensurable outcomes. Narrative synthesis is appropriate when the analytical task is to interrogate assumptions, reconcile conflicting evidence and evaluate the strength of inference across heterogeneous designs (Greenhalgh *et al*., 2018; Sukhera, 2022).

The choice of a narrative design does not license opacity. The search sources, search concepts, eligibility rules, appraisal criteria and synthesis logic are stated below so that the basis of each judgement can be examined. Reporting was guided by established quality criteria for narrative reviews, which emphasise justification of importance, statement of aims, description of literature searching, referencing of claims, evidence-calibrated language and appropriate presentation of data (Baethge *et al*., 2019).

### 2.2 Information Sources

Searching was conducted in Europe PMC, Crossref, Google Scholar, the Directory of Open Access Journals and Semantic Scholar. Trial status and design information were obtained from the ClinicalTrials.gov registry. Searching was supplemented by backward citation searching from recent reviews and consensus documents, by targeted author-based searching for research groups with sustained output in this field, and by related-record retrieval around key articles. Records were also checked for corrections and expressions of concern.

### 2.3 Search Strategy and Search Period

Search concepts were built from four elements: the disorder and its historical synonyms; the molecular substrate; the intervention classes of interest; and the outcome and methodological terms required to locate critical rather than descriptive literature. Representative strings included the following combinations, adapted to the syntax of each source: (“osteogenesis imperfecta” OR “brittle bone disease” OR “fragilitas ossium”) AND (*COL1A1* OR *COL1A2* OR “type I collagen” OR nosology OR classification OR “genotype-phenotype”); (“osteogenesis imperfecta”) AND (bisphosphonate* OR zoledronic OR denosumab OR teriparatide OR sclerostin OR setrusumab OR “transforming growth factor beta”); (“osteogenesis imperfecta”) AND (“gene therapy” OR CRISPR OR “base editing” OR “prime editing” OR “allele-specific” OR antisense OR “mesenchymal stem cell” OR “adeno-associated virus”); and (“osteogenesis imperfecta”) AND (mineralisation OR “bone material” OR “bone quality” OR “outcome measures” OR “natural history” OR penetrance).

The search period ran from 1 January 1979 to 15 June 2026. The starting year corresponds to the publication of the clinical classification that established the modern conceptual framework for the disorder and from which subsequent nosological and

therapeutic literature descends. Foundational publications predating this boundary were consulted where they were necessary to explain historical development, and are cited only for that purpose.

### 2.4 Eligibility Criteria

Peer-reviewed original research, systematic and narrative reviews, consensus statements, standard-setting documents and registered trial records were eligible. Eligible subject matter comprised the epidemiology, classification, molecular pathology, extraskeletal manifestations, outcome measurement, pharmacological management and cell or gene-directed treatment of osteogenesis imperfecta, together with methodological literature bearing directly on the appraisal of that evidence. Human studies of any design were eligible, as were animal and cell-based studies when they informed mechanism or preclinical therapeutic development. Publications in English were included.

Excluded material comprised single case reports without mechanistic or nosological novelty, conference abstracts, commercial and promotional documents, non-scholarly commentary, and any source whose bibliographic identity could not be confirmed. Preprints were excluded where peer-reviewed evidence on the same question was available. Studies of general osteoporosis, of other skeletal dysplasias and of monogenic low bone mass without a recognised link to this disorder were excluded unless cited for explicit comparison.

### 2.5 Screening, Duplicate Handling and Study Selection

Titles and abstracts were screened against the eligibility criteria, and records passing that stage were assessed in full text or, where full text was not accessible, through the complete abstract together with bibliographic metadata; sources that could not be verified in this way were not cited. Duplicates arising across sources were identified by digital object identifier, and where identifiers were absent, by matching of title, first author and publication year. Preference was given to the peer-reviewed version of record when a preprint and a published article described the same work. No numerical screening counts are reported, because selection was purposive and iterative rather than systematic, and reporting counts would imply a completeness that the design does not claim.

### 2.6 Critical Appraisal and Evidence Synthesis

Each included study was appraised for design adequacy relative to its question, cohort ascertainment and representativeness, presence and appropriateness of comparators, endpoint validity, sample size relative to the effect claimed, handling of confounding, duration of follow-up, and the distance between the observed finding and the clinical claim attached to it. Preclinical studies were additionally appraised for model fidelity to human disease, whether the intervention was delivered systemically or to isolated cells, and whether functional skeletal outcomes rather than molecular readouts were reported. Formal risk-of-bias instruments were not applied, because the included evidence spans designs for which no single instrument is valid and because several sources did not report the information such instruments require.

Themes were developed inductively from recurring points of tension rather than imposed in advance, and were retained when they were supported across independent research groups and when they carried implications for interpretation or practice. Where findings conflicted, the disagreement was examined for its probable source, distinguishing differences in cohort ascertainment, endpoint definition, treatment exposure and analytical approach from genuine biological inconsistency. Conclusions were graded qualitatively according to the convergence of independent evidence, the proximity of the endpoint to patient-relevant outcomes, and the extent to which alternative explanations had been excluded.

## 3. Historical Foundations and the Unsettled Architecture of Classification

### 3.1 From Clinical Syndromology to a Genetic Framework

The nineteenth-century literature on congenital bone fragility accumulated as a series of anatomical and clinical descriptions, and the eponyms attached to that literature persisted well into the twentieth century without conveying mechanistic information. The tradition was descriptive in the strict sense: patients were grouped by what could be observed at the bedside, on radiographs and at post-mortem examination. The persistence of these eponyms into contemporary clinical usage reflects the durability of descriptive categories rather than their explanatory power (Chetty *et al*., 2021).

The classification proposed by Sillence and colleagues represented a different order of claim. By combining inheritance pattern with radiographic severity, scleral colour and dental involvement, it asserted that the clinical continuum concealed discrete genetic entities and that these entities could be distinguished clinically (Sillence *et al*., 1979). The proposal proved productive precisely because it was falsifiable. It generated expectations about recurrence risk and prognosis that could be tested, and it provided the stratification framework within which later molecular work was conducted. Its practical value in the pre-molecular era is not in dispute.

Two features of the original scheme have since caused difficulty. The first is that severity, inheritance and specific clinical signs were treated as jointly informative when they are in fact only partly correlated. Blue sclerae occur across types and in unaffected individuals, and dentinogenesis imperfecta varies within families carrying identical variants. The second is that the numbered types were ordinal in appearance but not in severity, since type II denotes perinatal lethality while type III denotes the most severe survivable form. Clinicians and investigators have continued to use the numbering as though it were a severity scale, and the resulting ambiguity has propagated into trial eligibility criteria and cohort descriptions.

### 3.2 Genetic Expansion and the Collapse of Type-to-gene Correspondence

Identification of causative variants in the genes encoding the α1 and α2 chains of type I collagen accounted for the dominantly inherited majority of cases, and the discovery of recessive disease attributable to components of the collagen prolyl 3-hydroxylation complex demonstrated that fragility could arise without any defect in the collagen genes themselves (Cabral *et al*., 2007; Chang *et al*., 2010). Subsequent work extended the causative list to chaperones, secretory machinery, transcription factors, ion channels and secreted signalling molecules (Jovanovic *et al*., 2022; Jovanovic & Marini, 2024).

The consequence for classification was more disruptive than is usually acknowledged. As each new gene was described, it was assigned the next available Roman numeral, so that the numbering came to encode chronology of discovery rather than mechanism, severity or inheritance. A clinician told that a patient has type XX or type XXII learns nothing about the expected phenotype without consulting the underlying gene, and the numeral therefore performs no clinical work. Recessive entities defined by single families, such as those attributed to *MESD* and *CCDC134*, illustrate the problem clearly: the phenotypes are severe and biologically informative, but the numerals assigned to them convey neither the severity nor the mechanism (Ghosh *et al*., 2023; Ning *et al*., 2025).

An additional difficulty is that clinical severity is not partitioned by gene. Variants in the same gene produce phenotypes spanning near-normality to lethality, while variants in different genes converge on similar clinical presentations. Analyses of large clinically ascertained cohorts confirm broad within-gene variability and show that molecular category alone is a weak predictor of individual outcome (Garibaldi *et al*., 2022; Storoni *et al*., 2023). Any classification that assumes a stable mapping between gene and phenotype therefore encodes an inaccuracy.

### 3.3 Competing Frameworks and the Absence of Validation

Three responses to this situation now coexist in the literature. The first retains a clinical and radiographic grouping while annotating each group with its associated genes, an approach adopted in the international nosology of genetic skeletal disorders, which places the disorder within a broader taxonomy of bone fragility and dysplasia (Unger *et al*., 2023). The second proposes an explicitly dyadic designation in which a clinical descriptor is paired with the molecular cause, thereby preserving clinical communication while removing the pretence that a numeral encodes mechanism (Sillence, 2024). The third would reorganise the disorder around pathogenic mechanism, grouping entities by the cellular process disrupted rather than by clinical appearance (Yu *et al*., 2023). The mechanistic proposal has attracted substantive objection on the grounds that mechanistic assignment is often provisional, that several genes act in more than one pathway, and that replacing one unvalidated numbering system with another risks compounding rather than resolving confusion (Dalgleish *et al*., 2024).

The disagreement is genuine and the arguments on each side are coherent, which is itself diagnostic of the underlying problem. No framework has been evaluated prospectively against the outcomes that classification is supposed to serve, namely accuracy of prognosis, appropriateness of surveillance, selection of therapy and stratification of trial populations. Proposals are assessed against internal criteria such as parsimony, biological plausibility and consistency with existing usage, none of which is an empirical test. Discussions of the relative merits of classification for prophylaxis and genetic counselling remain largely argumentative rather than data-driven (Panzaru *et al*., 2023).

The practical cost is not merely terminological. Trial eligibility is commonly specified by clinical type, which means that populations described identically across studies may differ substantially in molecular composition and expected trajectory, complicating any attempt at cross-trial comparison. Where trials do specify a molecular criterion, they typically require a pathogenic variant in one of the collagen genes without further stratification, which merges quantitative and structural mechanisms that differ in bone material behaviour. Until classification is tested against outcomes, its principal defensible function is communication, and claims that any current scheme improves clinical decision-making should be treated as untested.

Table 1 compares the principal frameworks in use, the basis on which each is constructed, the purposes each serves adequately and the specific respects in which each fails. The comparison shows that the frameworks differ less in accuracy than in what they choose to encode, and that none currently discharges the prognostic function most often attributed to classification. The table also makes visible a shared weakness: every scheme presupposes that assignment is stable, whereas variant reclassification and evidence of incomplete penetrance both indicate that assignment can change over time.

## 4. Molecular Pathology: Collagen, the Wider Osteoblast Network and Bone Material Failure

### 4.1 Quantitative and Structural Defects of Type I Collagen

Type I collagen is a heterotrimer of two α1 chains and one α2 chain assembled into a triple helix whose stability depends on glycine occupying every third position of the helical domain. Two mechanistically distinct classes of defect follow from this architecture. Variants that abolish expression from one *COL1A1* allele reduce the quantity of structurally normal collagen, producing haploinsufficiency and, characteristically, a milder phenotype. Variants that substitute a bulkier residue for a helical glycine, or that disturb splicing in a manner preserving the reading frame, allow an abnormal chain to be incorporated into the trimer, where it exerts a dominant-negative effect on folding, secretion and matrix assembly (Forlino & Marini, 2016; Marini *et al*., 2017).

**Table 1. Comparison of classification frameworks for osteogenesis imperfecta, with the basis, function and principal limitation of each**

| Framework | Organising basis | Function served adequately | Principal limitation | Validated against clinical outcomes | Supporting references |
|---|---|---|---|---|---|
| Descriptive and eponymous tradition | Clinical and anatomical observation | Historical continuity; recognition of the phenotypic spectrum | Conveys no mechanistic or prognostic information; eponyms are inconsistently applied | No | Chetty *et al*. (2021) |
| Original four-type clinical classification | Inheritance pattern, radiographic severity, scleral hue, dental involvement | Bedside communication; recurrence-risk counselling in the pre-molecular era | Numbering is not ordinal by severity; component signs are only partly correlated with type | No | Sillence *et al*. (1979); Chetty *et al*. (2021) |
| Sequentially numbered genetic types | Chronological order of gene discovery | Indexing of newly described entities in the literature | Numerals encode discovery order rather than mechanism or severity; several types are defined by very few families | No | Ghosh *et al*. (2023); Jovanovic & Marini (2024); Ning *et al*. (2025) |
| International skeletal dysplasia nosology grouping | Clinical and radiographic grouping annotated with associated genes | Placement within the wider taxonomy of skeletal disorders; standardised terminology | Retains clinical grouping whose boundaries are indistinct at the mild and severe extremes | No | Unger *et al*. (2023) |
| Dyadic clinical-plus-molecular designation | Explicit pairing of a clinical descriptor with the causative gene | Preserves clinical communication without implying that a numeral encodes mechanism | Requires molecular confirmation, which is unavailable in many settings | No | Sillence (2024) |
| Mechanism-based reorganisation | Disrupted cellular process | Alignment of nomenclature with biology; potential relevance to mechanism-directed therapy | Mechanistic assignment is provisional for several genes, and some genes act in multiple pathways | No | Dalgleish *et al*. (2024); Yu *et al*. (2023) |

*Note. The final column records whether the framework has been evaluated prospectively against prognostic accuracy, surveillance appropriateness, therapeutic selection or trial stratification. Absence of validation is not evidence that a framework is inaccurate; it indicates that the claim of clinical utility remains untested*

This distinction is the most robust genotype-phenotype relationship in the field, and it has direct therapeutic implications, since haploinsufficiency is in principle addressable by increasing output from the intact allele, whereas dominant-negative disease requires selective removal or correction of the mutant product. The distinction is nevertheless coarser than it is often presented. The severity of structural variants depends on the position of the substitution along the helix, on the identity of the substituting residue, and on whether the affected chain is α1 or α2, and the resulting severity gradients are irregular rather than smoothly graded. Structural rearrangements affecting one or both collagen genes add a further layer of complexity that conventional variant classification handles poorly, since deletions and duplications may produce either quantitative or structural consequences depending on their boundaries (Batkovskyte *et al*., 2025).

### 4.2 The Limits of Genotype-based Prediction

Cohort studies consistently report substantial phenotypic variability among individuals carrying identical variants, including within families, which indicates that variant identity alone cannot determine outcome (Garibaldi *et al*., 2022). Analyses of large clinically ascertained series relate molecular category to severity distributions but not to individual prediction (Storoni *et al*., 2023). In adults, prospective follow-up indicates that low areal bone mineral density and particular variant classes each contribute to fracture probability, yet the confidence intervals around the reported associations are wide and the cohorts small, so the practical predictive value for an individual remains modest (Blandin *et al*., 2025).

A more fundamental challenge has come from population-scale sequencing. Analysis of genome sequencing data from approximately half a million adults in a national biobank identified pathogenic and likely pathogenic collagen gene variants in a small number of participants whose mean age at recruitment exceeded fifty years, and found overall penetrance substantially below expectation, at approximately forty-one per cent for *COL1A1* and twenty-one per cent for *COL1A2*; penetrance rose to about seventy-three per cent when only truncating *COL1A1* variants were considered (Pagnamenta *et al*., 2026). The finding requires careful interpretation, because a cohort of healthy volunteers recruited in middle age is depleted of severe early-onset disease and because phenotype ascertainment relied on diagnostic coding, self-report and heel densitometry rather than on skeletal assessment. The direction of the result is nevertheless important. Penetrance estimates derived from referral cohorts are inflated by the ascertainment process that assembled them, and estimates intended to support genomic newborn screening cannot be taken from clinic populations without adjustment. Independent replication in cohorts with better skeletal phenotyping is required before the magnitude of the effect can be regarded as established.

### 4.3 Post-translational Modification, Folding and Secretion

Recessive disease demonstrated that fragility can arise from defects in the machinery that modifies and handles collagen rather than in collagen itself. Deficiency of prolyl 3-hydroxylase 1 produces a severe recessive phenotype resembling the lethal and severe dominant forms, and the enzyme functions within a complex whose components are mutually stabilising, so that loss of any one member destabilises the others (Cabral *et al*., 2007; Chang *et al*., 2010). This finding established a general principle that has since been extended: the pathway from transcript to secreted, correctly modified trimer contains numerous points at which failure produces a comparable clinical result.

Later work has added chaperone and secretory components to this account. Variants in the endoplasmic reticulum protein *MESD* cause cellular stress and intracellular aggregation of type I collagen, linking a folding defect to a recognisable severe phenotype (Ghosh *et al*., 2023). A rare untranslated region variant in *SEC24D*, a component of the coat complex responsible for export from the endoplasmic reticulum, has been shown to impair translation, and the same work outlines a route to RNA-level rescue (Essawi *et al*., 2026). Loss of the trimeric intracellular cation channel TRIC-B dysregulates calcium-dependent signalling and, through inhibition of calcium/calmodulin-dependent protein kinase II, alters SMAD-mediated pathways in osteoblasts (Besio *et al*., 2023).

The common inference drawn from these studies, that heterogeneous upstream defects converge on a shared final pathway of impaired collagen output and disturbed osteoblast function, is plausible and is supported by the phenotypic similarity of the resulting conditions. The inference should nevertheless be held loosely. Convergence has been asserted more often than it has been demonstrated, since the studies establishing each mechanism use different models, different readouts and different endpoints, and direct comparisons across mechanisms within a single experimental system are rare. Several mechanisms rest on findings from one or two families supported by cell-based functional work, which establishes causation for the variant but does not establish that the proposed mechanism accounts for the full phenotype.

### 4.4 Signalling Defects and the Special Case of Type V Disease

A distinct group of entities arises not from collagen handling but from disturbance of osteoblast differentiation and bone formation signalling. The recurrent variant in the 5′ untranslated region of *IFITM5* that causes type V disease creates a novel start codon and produces a phenotype with hyperplastic callus, interosseous membrane calcification and radial head dislocation that is unlike collagen-related disease. Work in an inducible mouse model carrying the human variant confirms that the mutation drives aberrant osteogenesis rather than simply reducing collagen output (Tan *et al*., 2025). Complementary studies indicate that the variant acts through an ERK- and SOX9-dependent defect in osteoprogenitor differentiation, which offers a coherent explanation for the distinctive tissue-level features (Marom *et al*., 2024).

X-linked bone fragility attributable to *PLS3* occupies a contested position at the boundary of the disorder, since affected individuals present with early-onset osteoporosis and fractures without the extraskeletal features characteristic of collagen-related disease. Cohort description indicates a phenotype that overlaps clinically but differs in its natural history and in the pattern of skeletal

involvement (Zervou *et al*., 2025). Whether such conditions are best regarded as forms of osteogenesis imperfecta or as distinct monogenic osteoporoses is a question that classification frameworks answer differently, and the answer has practical consequences for trial eligibility and for the interpretation of cohort prevalence.

### 4.5 Bone Material Quality and the Paradox of Increased Mineral Density

The most consequential mechanistic insight for clinical interpretation concerns bone material rather than bone mass. Quantitative backscattered electron imaging of iliac bone from affected individuals shows a matrix that is more highly mineralised than normal, with altered mineral particle characteristics and abnormal mineralisation kinetics, and these features persist across disease types and are modified but not normalised by antiresorptive treatment (Misof & Fratzl-Zelman, 2024). Mineralisation kinetics are accelerated even in the mildest quantitative form of the disease, indicating that the abnormality is intrinsic to the matrix rather than secondary to high turnover (Misof *et al*., 2023).

**Table 2. Principal molecular mechanism groups in osteogenesis imperfecta, with representative genes, attributed cellular consequence and residual uncertainty**

| Mechanism group | Representative genes | Attributed cellular consequence | Usual inheritance | Principal residual uncertainty | Supporting references |
|---|---|---|---|---|---|
| Quantitative collagen deficiency | *COL1A1* | Reduced output of structurally normal type I collagen | Autosomal dominant | Penetrance in unselected populations is lower than clinic-based estimates imply | Forlino & Marini (2016); Pagnamenta *et al*. (2026) |
| Structural collagen defect (dominant negative) | *COL1A1*, *COL1A2* | Incorporation of abnormal chains, impaired folding and secretion, disordered matrix assembly | Autosomal dominant | Severity gradients by helical position and residue identity are irregular and incompletely predictive | Batkovskyte *et al*. (2025); Marini *et al*. (2017) |
| Defective collagen prolyl 3-hydroxylation | *CRTAP*, *P3H1*, *PPIB* | Loss of a mutually stabilising complex; overmodification and delayed folding | Autosomal recessive | Relative contribution of overmodification versus reduced secretion to tissue-level failure | Cabral *et al*. (2007); Chang *et al*. (2010) |
| Chaperone, folding and export failure | *SERPINH1*, *FKBP10*, *MESD*, *SEC24D* | Endoplasmic reticulum stress, intracellular collagen aggregation, impaired export | Autosomal recessive | Whether convergence on a shared final pathway is genuine or inferred from phenotypic similarity | Essawi *et al*. (2026); Ghosh *et al*. (2023) |
| Ion homeostasis and intracellular signalling | *TMEM38B* (TRIC-B) | Dysregulated calcium-dependent signalling with downstream SMAD pathway alteration | Autosomal recessive | Extent to which signalling change, rather than collagen output, determines the skeletal phenotype | Besio *et al*. (2023) |
| Osteoblast differentiation and bone formation signalling | *IFITM5*, *SERPINF1*, *WNT1*, *SP7*, *CCDC134* | Impaired osteoprogenitor differentiation and defective bone formation | Autosomal dominant (*IFITM5*) or recessive | Several assignments rest on single families with supporting cell-based work only | Marom *et al*. (2024); Ning *et al*. (2025); Tan *et al*. (2025) |
| X-linked bone fragility at the diagnostic boundary | *PLS3* | Altered osteocyte and osteoblast mechanotransduction | X-linked | Whether the condition belongs within this disorder or among monogenic osteoporoses | Zervou *et al*. (2025) |
| Matrix material abnormality (common downstream feature) | Applies across groups | Hypermineralised matrix with altered mineral particle characteristics and accelerated mineralisation kinetics | Not applicable | Quantitative relationship between material abnormality and individual fracture risk | Misof & Fratzl-Zelman (2024); Misof *et al*. (2023) |

*Note. Gene symbols are italicised in accordance with standard gene nomenclature; protein names are not italicised. Mechanism groups are analytical categories used for synthesis and do not correspond to numbered clinical types. The final row describes a tissue-level feature observed across mechanism groups rather than a separate genetic category*

The clinical implication is direct and frequently understated. Bone that is more densely mineralised but composed of defective collagen scaffolding is stiffer and more brittle, so gains in mineral density need not translate into gains in fracture resistance and may in some circumstances be neutral or counterproductive. Areal bone mineral density measured by dual-energy X-ray absorptiometry, which remains the dominant efficacy endpoint in trials, is therefore measuring a quantity whose relationship to the

outcome of interest is uncertain in this specific disorder. This is not a peripheral methodological observation; it provides the most economical explanation for the recurring pattern in which interventions produce substantial densitometric change and much smaller or undetectable fracture effects.

Table 2 summarises the principal mechanistic groups, their representative genes, the cellular consequence attributed to each and the specific uncertainty attaching to that attribution. The table is organised by mechanism rather than by numbered type in order to make visible both the convergence claimed in the literature and the unevenness of the evidence supporting each assignment. Comparison across rows shows that the strength of mechanistic evidence varies by an order of magnitude, from mechanisms characterised in multiple independent models to those supported by a single family and one cell-based assay.

## 5. The Multisystem Phenotype and the Problem of Outcome Measurement

### 5.1 Extraskeletal Involvement as a Systematic Rather than Incidental Feature

Because type I collagen is the principal structural protein of skin, tendon, ligament, sclera, dentine, the cardiovascular adventitia and the pulmonary interstitium, extraskeletal involvement should be expected rather than treated as a complication. Evidence accumulated over the past decade supports this expectation across several organ systems.

Hearing loss is among the best characterised extraskeletal features. A Danish nationwide register-based cohort study comparing individuals with the disorder against a matched reference population found a substantially elevated subhazard for any hearing loss and a markedly elevated subhazard for otosclerosis or middle ear surgery, with onset at a younger median age than in the reference population and prevalence rising with age (Haumann *et al*., 2024). Register-based designs of this kind have well-recognised limitations, including dependence on diagnostic coding and inability to characterise audiometric severity, yet the magnitude and consistency of the effect make misclassification an implausible complete explanation.

Pulmonary involvement has historically been attributed to chest wall deformity and scoliosis, which would make it a mechanical consequence of skeletal disease. Detailed assessment of lung function and structure in children and young adults indicates abnormalities consistent with an intrinsic parenchymal contribution in addition to extrinsic restriction (Gochuico *et al*., 2023). The distinction matters therapeutically, because intrinsic lung abnormality would not be expected to respond to correction of spinal deformity. The study population was drawn from a specialist referral setting and was modest in size, so the relative contribution of intrinsic and extrinsic mechanisms across the wider population remains unquantified.

Cardiovascular involvement is the least well evidenced of the major extraskeletal domains. An international expert group used a modified Delphi process to describe clinical characteristics, generate care recommendations and identify research priorities, an approach that is appropriate when primary evidence is sparse but which produces consensus opinion rather than empirical estimates (Folkestad *et al*., 2025). Recommendations derived in this way should be implemented as pragmatic guidance and cited as expert opinion, not as demonstrated benefit.

Dental involvement extends beyond dentinogenesis imperfecta. Systematic review with meta-analysis of prevalence and comparative studies reports pulp obliteration, dental impaction and tooth agenesis as frequent findings, with impaction substantially more common than in unaffected comparison groups and tooth discolouration more frequent in the more severe clinical types (Prado *et al*., 2023). The included studies were predominantly observational with heterogeneous diagnostic criteria, so pooled prevalence estimates carry wide uncertainty even where the direction of effect is consistent.

### 5.2 Natural History, Registries and the Evidentiary Base for Anticipatory Care

Systematic collation of the natural history literature demonstrates that signs, symptoms and events emerge across the lifespan in a partly predictable sequence, with deformity and scoliosis progressing during childhood and cardiac, ocular, auditory and joint problems emerging in adulthood, generally earlier and more frequently than in the general population (Gatti *et al*., 2026). The same synthesis records wide variation in age at diagnosis, with severe presentations identified in early childhood and milder presentations diagnosed at highly variable ages. This variability is not merely descriptive; it determines the composition of any cohort assembled through clinical services and therefore biases the natural history estimates derived from such cohorts.

Registry and administrative data supply population-level estimates that referral cohorts cannot. National registry linkage in the Netherlands quantifies both incidence and healthcare utilisation and shows that the excess hospitalisation burden is concentrated in childhood (Storoni *et al*., 2022). National hospital discharge data in France demonstrate that mortality is heavily concentrated in the neonatal period, with stillbirth accounting for the majority of recorded deaths (Philippoteaux *et al*., 2025). Administrative datasets of this kind are prone to diagnostic misclassification and cannot distinguish clinical types reliably, but they are considerably less susceptible to referral bias than specialist cohorts and provide the more defensible basis for estimating population burden.

The humanistic burden has been characterised through structured survey work in adults, which documents impact on daily function, employment and psychological wellbeing extending well beyond fracture events (van Welzenis *et al*., 2024). Surveys recruited through patient organisations over-represent those who are engaged with advocacy networks and probably under-represent both the mildest and the most severely disabled, so the reported distributions should be read as indicative rather than as population estimates.

### 5.3 Standardised Outcome Sets and the Surrogate Endpoint Problem

International standard sets of outcome measures now exist for comprehensive assessment, for hearing and for oral health, developed through structured consensus involving clinicians and patient representatives (Blokland *et al*., 2024; Goderie *et al*., 2023; Nijhuis *et al*., 2021). Their existence addresses a real deficiency, since heterogeneous outcome reporting has historically prevented meaningful comparison across studies. Their impact to date is limited, because adoption in interventional trials has been slow and because the sets describe what should be measured without resolving which measures are responsive enough to detect treatment effects within feasible trial durations.

The deeper problem is the continued reliance on areal bone mineral density as the primary efficacy endpoint. The mechanistic argument developed in Section 4.5 predicts that densitometric gain will overstate mechanical benefit in a disorder characterised by hypermineralised and abnormally organised matrix, and the empirical pattern across the therapeutic literature is consistent with that prediction. Fracture is the outcome that matters to patients, yet fracture counts are noisy, depend on activity level and ascertainment, decline with age in the mildest forms and require either large samples or long follow-up to yield stable estimates. In a rare disorder with a wide severity range, these requirements are difficult to satisfy, which explains why densitometric surrogates have persisted despite their known limitations. No surrogate endpoint in this field has been formally qualified against fracture outcomes, and until such qualification is attempted the interpretation of densitometric trial results will remain conditional.

Table 3 sets out the principal outcome domains, the measures currently used, the specific limitation of each measure in this disorder and the resulting implication for trial interpretation. Reading across the table shows that the measures that are most responsive are the least closely tied to patient-relevant outcomes, and that the measures most closely tied to those outcomes are the least statistically tractable. That inverse relationship, rather than any single deficient instrument, is the central methodological obstacle in the field.

**Table 3. Outcome domains and measures used in osteogenesis imperfecta research, with domain-specific limitations and implications for interpretation**

| Outcome domain | Measures commonly used | Principal limitation in this disorder | Implication for trial interpretation | Supporting references |
|---|---|---|---|---|
| Bone mass | Areal bone mineral density by dual-energy X-ray absorptiometry; Z-scores | Measures mineral quantity in a matrix that is already abnormally mineralised; short stature and deformity distort areal measurement | Densitometric gain cannot be assumed to indicate improved fracture resistance | Misof & Fratzl-Zelman (2024); Misof *et al*. (2023) |
| Bone material and microarchitecture | Quantitative backscattered electron imaging; high-resolution peripheral quantitative computed tomography; finite element estimates of strength | Requires specialist facilities; reference data in children are limited; relationship to clinical fracture not quantified | Provides mechanistically relevant information but is not yet a qualified surrogate | Glorieux *et al*. (2024); Misof & Fratzl-Zelman (2024) |
| Fracture | Annualised clinical fracture rate; radiographic vertebral morphometry | High variance; dependent on activity and ascertainment; requires large samples or long follow-up in a rare disorder | The only patient-relevant skeletal endpoint, but frequently underpowered | Dwan *et al*. (2016); Wang *et al*. (2025) |
| Growth and deformity | Height and length Z-scores; scoliosis angle; long bone deformity indices | Confounded by surgical intervention and by secular changes in management | Difficult to attribute change to pharmacological intervention | Gatti *et al*. (2026) |
| Extraskeletal function | Audiometry; pulmonary function testing; cardiac imaging | Standard sets exist but are unevenly adopted; several recommendations rest on expert consensus | Extraskeletal benefit of skeletal therapies is largely unmeasured | Folkestad *et al*. (2025); Gochuico *et al*. (2023); Goderie *et al*. (2023); Haumann *et al*. (2024) |
| Oral health | Standardised oral health and occlusion outcome set | Recently introduced; comparative data across cohorts remain limited | Dental outcomes are rarely reported in interventional studies | Blokland *et al*. (2024); Prado *et al*. (2023) |
| Patient-reported outcome and burden | Disease-specific and generic quality-of-life instruments; structured surveys | Recruitment through patient organisations introduces selection; instruments are variably validated in this population | Reported burden is indicative rather than representative | Nijhuis *et al*. (2021); van Welzenis *et al*. (2024) |

*Note. Dual-energy X-ray absorptiometry measures areal rather than volumetric density and is sensitive to bone size, which is systematically reduced in this population. Qualification of a surrogate endpoint requires demonstration that treatment effects on the surrogate predict treatment effects on the clinical outcome; no such qualification has been reported for any measure in this table*

## 6. Critical Appraisal of Pharmacological Management

### 6.1 Bisphosphonates: Reliable Densitometric Effect, Uncertain Fracture Effect

Cyclical intravenous pamidronate in severely affected children produced large increases in bone mineral density, reduced bone pain and apparent functional improvement in an uncontrolled observational study whose findings were rapidly incorporated into practice (Glorieux *et al*., 1998). The historical influence of that report is difficult to overstate, and it remains the point from which contemporary practice descends (Rauch & Glorieux, 2004). The design, however, provided no control group, and the improvement observed in growing children receiving intensive multidisciplinary care cannot be attributed to the drug with confidence.

Subsequent controlled evidence has confirmed the densitometric effect and left the fracture effect unresolved. Systematic review of randomised trials concludes that bisphosphonates consistently increase bone mineral density while the evidence for fracture reduction is inconsistent, with trials differing in agent, route, dose, duration and population (Dwan *et al*., 2016). A network meta-analysis of nine randomised trials including 595 children reinforces this asymmetry: at two years all active treatments improved lumbar spine areal bone mineral density relative to placebo, with the largest point estimate for pamidronate, and zoledronic acid produced the most favourable lumbar spine Z-score changes, yet in the limited fracture analysis only olpadronate reduced total fracture numbers relative to placebo, and the certainty of evidence for the densitometric comparisons was rated as low to moderate (Wang *et al*., 2025). The pattern is precisely that predicted by the bone material argument: the surrogate responds reliably, and the clinical outcome does not.

Several further limitations deserve emphasis. Trials have been short relative to the duration of exposure in clinical practice, where treatment may continue for many years in growing children. Delayed osteotomy healing during pamidronate treatment has been documented, although fracture healing was not similarly delayed, which introduces a specific interaction with surgical management that dosing schedules must accommodate (Munns *et al*., 2004). Long-term skeletal consequences of prolonged suppression of remodelling during growth remain incompletely characterised, and the optimal duration, cessation criteria and management after discontinuation are not established by controlled evidence.

### 6.2 Denosumab and Sequential Regimens

Denosumab attracted interest as a reversible antiresorptive alternative, particularly for recessive forms in which bisphosphonate response appeared attenuated. The first prospective comparative study randomised 84 children and adolescents to denosumab or a single infusion of zoledronic acid and found substantial increases in bone mineral density at lumbar spine, femoral neck and total hip in both groups over twelve months, together with improvement in vertebral height and projected area (J. Liu *et al*., 2024). The safety finding was more consequential than the efficacy finding: rebound hypercalcaemia was common, and a substantial minority of affected participants reached hypercalcaemic crisis, a problem mitigated by switching to zoledronic acid.

This result illustrates a broader point about the transfer of adult osteoporosis pharmacology to a paediatric genetic disorder. Denosumab's reversibility, which is advantageous in adults, becomes hazardous in growing children with high baseline remodelling, and the dosing interval derived from adult practice appears inappropriate. The study was open in design, of twelve months' duration and powered for densitometric rather than fracture outcomes, so its principal contribution is the characterisation of a serious and predictable adverse effect rather than the demonstration of comparative efficacy.

Sequential anabolic-then-antiresorptive strategies represent a rational response to the limitations of antiresorptive monotherapy in adults, and a randomised trial of teriparatide followed by zoledronic acid against standard care has been designed with fracture reduction as its target (Hald *et al*., 2023). Trials of this design address the endpoint problem directly by accepting the sample size and duration that a fracture endpoint requires. Until their results are available, adult management continues to rest largely on extrapolation from osteoporosis practice, and the evidence base specific to adults with this disorder remains thin (W. Liu *et al*., 2024).

### 6.3 Sclerostin Inhibition: The Clearest Illustration of the Surrogate Problem

Sclerostin is a negative regulator of WNT-mediated bone formation, and its inhibition offers an anabolic mechanism that is conceptually well matched to a disorder of deficient bone formation. Preclinical support is genuine but qualified: genetic inactivation of sclerostin in a mouse model of severe dominant disease improved bone mass, although the correction was incomplete and did not normalise the material properties of the matrix (Marulanda *et al*., 2023). This limitation was visible before clinical development advanced and was consistent with the expectation that adding more of an intrinsically defective matrix cannot restore normal mechanical competence.

The phase 2b evaluation of setrusumab in adults randomised 110 participants with a clinical diagnosis and a confirmed collagen gene variant to three dose levels or placebo over twelve months. The primary endpoint, change in distal radial trabecular volumetric bone mineral density, was not met, while microstructural finite element estimates of failure load and stiffness improved significantly at the higher doses (Glorieux *et al*., 2024). Parallel investigation of bone material properties in treated adults found no adverse effect on matrix characteristics, which addressed a legitimate mechanistic concern about anabolic therapy in an already hypermineralised matrix (Rummler *et al*., 2024). Taken together, these results supported progression to definitive evaluation while leaving the relationship between strength estimates and clinical fracture unestablished.

The subsequent phase 3 programme is registered with a primary endpoint of annualised clinical fracture rate, and the registry record indicates that primary completion occurred in October 2025 (U.S. National Library of Medicine, 2026). At the search date for this review, no peer-reviewed report of the phase 3 outcome had been identified in the accessible literature, and the fracture efficacy of sclerostin inhibition in this disorder therefore cannot be regarded as established or refuted on the basis of published evidence. This situation is itself instructive. A development programme built on densitometric and microstructural surrogates advanced to definitive testing before the predictive validity of those surrogates had been demonstrated, and the field consequently awaits a single trial to resolve a question that a qualified surrogate would have addressed earlier and at lower cost.

### 6.4 Transforming Growth Factor Beta Inhibition and Other Pathway-Directed Approaches

Excess transforming growth factor beta signalling has been implicated in the pathogenesis of both dominant and recessive forms, and inhibition of the pathway improves skeletal phenotype in animal models. A first-in-disease clinical evaluation of a neutralising antibody in adults reported dose-dependent increases in lumbar spine bone mineral density, with the response differing by disease type in a manner suggesting that mechanism-stratified selection may be necessary (Song *et al*., 2022). The study was small and open in design, and its endpoints were densitometric, so it should be read as establishing feasibility and generating hypotheses rather than as demonstrating benefit. Reviews of the developing pharmacological landscape in adults reach similar conclusions, describing several plausible mechanisms with limited controlled evidence behind any of them (W. Liu *et al*., 2024).

**Table 4. Pharmacological approaches in osteogenesis imperfecta, with the strongest available evidence, endpoints assessed and principal interpretive limitation**

| Approach | Strongest available evidence | Effect on bone mineral density | Evidence on fracture | Principal limitation | Supporting references |
|---|---|---|---|---|---|
| Cyclical intravenous pamidronate | Uncontrolled observational study in severely affected children; later randomised trials | Large and consistent increase | Not demonstrated in the original uncontrolled study | Absence of a control group in the foundational report; short trial durations thereafter | Dwan *et al*. (2016); Glorieux *et al*. (1998) |
| Bisphosphonates as a class | Systematic review of randomised trials; network meta-analysis of nine randomised trials in 595 children | Consistent increase at two years across agents | Inconsistent; benefit shown for one agent only in a limited analysis | Certainty of evidence rated low to moderate; heterogeneous agents, doses and durations | Dwan *et al*. (2016); Wang *et al*. (2025) |
| Zoledronic acid | Randomised comparison with denosumab; network meta-analysis | Favourable lumbar spine Z-score change relative to comparators | Not separately established | Comparative fracture data lacking; interaction with osteotomy healing requires scheduling care | J. Liu *et al*. (2024); Munns *et al*. (2004); Wang *et al*. (2025) |
| Denosumab | Prospective randomised comparison in 84 children and adolescents over 12 months | Substantial increase at spine, femoral neck and total hip | Not assessed as a primary endpoint | Rebound hypercalcaemia was common and sometimes severe; adult-derived dosing appears unsuitable | J. Liu *et al*. (2024) |
| Sequential teriparatide then zoledronic acid | Randomised trial designed with a fracture endpoint in adults | Not yet reported | Trial designed to address this directly | Results not yet available; adult practice currently relies on extrapolation | Hald *et al*. (2023); W. Liu *et al*. (2024) |
| Sclerostin inhibition | Phase 2b randomised trial in 110 adults; supporting bone material analysis; registered phase 3 programme | Primary volumetric density endpoint not met; estimated bone strength improved at higher doses | Phase 3 primary endpoint; no peer-reviewed report identified at the search date | Development advanced on unqualified surrogates; preclinical correction was incomplete | Glorieux *et al*. (2024); Marulanda *et al*. (2023); Rummler *et al*. (2024); U.S. National Library of Medicine (2026) |
| Transforming growth factor beta inhibition | Small open-label clinical evaluation with supporting animal work | Dose-dependent increase at lumbar spine, differing by disease type | Not assessed | Small, open design; densitometric endpoints; suggests need for mechanistic stratification | Song *et al*. (2022) |

*Note. Certainty ratings reported here are those assigned by the cited syntheses. Absence of demonstrated fracture benefit denotes absence of adequate evidence rather than demonstrated absence of effect*

The broader lesson from the pathway-directed literature is that mechanistic stratification has been proposed repeatedly and implemented rarely. Trials continue to enrol populations defined by clinical type or by the presence of any collagen gene variant, despite mechanistic arguments that quantitative and structural defects, and dominant and recessive forms, should respond differently. Enrolling mechanistically heterogeneous populations dilutes true effects and complicates interpretation of null results, and this design choice may account for part of the discrepancy between preclinical promise and clinical performance.

Table 4 compares the principal pharmacological classes across the evidence available for each, the endpoints on which that evidence rests and the specific limitation constraining interpretation. The comparison shows that the strength of evidence for densitometric effect is consistently greater than the strength of evidence for fracture effect across every class, and that safety signals rather than efficacy findings have been the most decisive clinical outputs of several recent trials.

## 7. Cell and Gene-Directed Therapy: Evidence, Mechanistic Rationale and Translational Barriers

### 7.1 Why the Dominant Mechanism Defines the Therapeutic Problem

Any curative strategy must contend with the fact that most disease is dominant negative rather than loss of function. Supplying additional normal collagen does not remove the abnormal chains that disrupt trimer folding and matrix assembly, so restoration of normal matrix requires selective suppression or correction of the mutant allele rather than augmentation of the normal one. The recessive forms, in which a functional protein is absent, present a comparatively conventional gene replacement problem. This asymmetry explains why preclinical demonstrations of benefit have been achieved more readily in recessive models, and why extrapolation from those demonstrations to the dominant-negative majority is not straightforward (Schindeler *et al*., 2022).

Three further constraints apply irrespective of mechanism. The target cell population comprises osteoblasts and their skeletal progenitors, which are dispersed throughout the skeleton and are not readily accessible to systemic vectors. The therapeutic window is developmental, because the deformity and fracture burden accumulate during rapid growth, and correction after skeletal maturity cannot reverse established deformity. Durability is demanding, because episomal vector genomes are diluted by cell division in a growing skeleton, so transient expression is unlikely to yield lasting benefit.

### 7.2 Mesenchymal Stromal Cell Transplantation

Transplantation of mesenchymal stromal cells has the longest clinical history of any cell-based approach in this disorder. The rationale combines engraftment of donor osteoprogenitors capable of secreting normal collagen with trophic effects on host bone cells, and early case reports described increases in total body bone mineral content, improved growth velocity and reduced fracture frequency (Götherström *et al*., 2021).

Systematic appraisal places this evidence in perspective. A review that screened forty studies identified nine human clinical reports, comprising case reports and case series and involving twelve patients in total; five reported increased total body bone mineral content and four reported reduced fracture rates after transplantation, with studies using bone marrow-derived cells rated higher in methodological quality than those using fetal cells (Khorasani *et al*., 2025). An evidence base of twelve patients across uncontrolled designs cannot support conclusions about efficacy. Donor cell engraftment in reported cases has been low and often transient, which sits uneasily with the durable clinical improvements described, and the alternative explanation of trophic or paracrine effect has not been tested against a control condition.

Prospective evaluation is proceeding cautiously. An exploratory open-label multicentre phase I/II trial has been designed to assess postnatal, or prenatal and postnatal, administration of allogeneic expanded fetal mesenchymal stem cells in infants and fetuses with severe disease, with safety as the principal focus (Sagar *et al*., 2024). A separate protocol describes multiple intravenous and intraosseous doses of fetal liver-derived cells in fifteen young children, again with safety and tolerability as the primary aim and fracture rate, bone mineral density, growth and biochemical turnover as secondary endpoints (Madhuri *et al*., 2025). Both are appropriately framed as exploratory. Neither is controlled, and both will therefore establish feasibility and safety rather than efficacy. Prenatal administration raises additional considerations, since a fetal intervention followed by continuing postnatal treatment complicates the attribution of any observed benefit to the prenatal component.

### 7.3 Gene Editing and Nucleic Acid-Directed Correction

Editing approaches divide into three strategies with different requirements. Allele-specific silencing removes the mutant transcript while leaving the normal allele intact, and is conceptually well matched to dominant-negative disease, but it requires reagents specific to each variant and most causative variants are private to individual families. Correction of the causative variant restores normal sequence and is durable, but conventional homology-directed repair is inefficient in post-mitotic and slowly dividing cells. Modulation of expression, whether by promoter editing or by RNA-level intervention, avoids double-strand breaks but achieves quantitative rather than qualitative correction.

Published work spans all three. Correction of a dominant cis-double-variant in *COL1A1* using CRISPR/Cas9 in patient-derived induced pluripotent stem cells increased osteogenic capacity, demonstrating that correction is achievable and functionally consequential at the cellular level (Cao *et al*., 2023). Editing of the mutated *COL1A1* gene in induced pluripotent stem cells has similarly been shown to restore osteogenesis in vitro (Jung *et al*., 2021). Base editing of the *COL1A1* promoter has been used to modulate collagen I expression in fibroblasts without introducing double-strand breaks, illustrating a route to quantitative adjustment that may suit haploinsufficiency (Daliri *et al*., 2025). At the RNA level, characterisation of a translational defect caused

by an untranslated region variant in *SEC24D* has been accompanied by a proposed rescue strategy, indicating that transcript-directed correction may be feasible for specific variant classes (Essawi *et al*., 2026).

In vivo demonstrations remain few but are more informative than cell-based work. Systemic administration of dual recombinant adeno-associated virus vectors delivering CRISPR/Cas9 together with a donor template to correct a *Col1a2* mutation in a mouse model of severe dominant disease reduced bone matrix turnover, altered osteoblast and osteoclast development and improved the osteocyte network (Yang *et al*., 2024). In a recessive context, intraosseous delivery of an adeno-associated virus serotype 9 vector encoding *Wnt1* in a newly generated rat model recapitulating recessive disease recovered bone density and mechanical strength (Li *et al*., 2026). These studies establish that skeletal delivery is achievable and that functional improvement follows, but the interventions were delivered to young animals with short skeletal growth periods, the follow-up was limited, and neither addresses the durability problem posed by prolonged human growth.

Broader reviews of gene editing for collagen disorders describe the technical landscape and the remaining obstacles in similar terms, emphasising delivery, allele discrimination and the regulatory challenge of variant-specific therapeutics (Kocsy *et al*., 2025; Schindeler *et al*., 2022). The most important observation for a critical reading is negative: no completed clinical trial of gene editing or gene transfer in this disorder was identified in the accessible literature at the search date, and descriptions of these approaches as near-term therapeutic options are therefore not supported by clinical evidence.

### 7.4 Translational Barriers that Mechanism Alone will not Overcome

Four barriers recur across the preclinical literature and are addressed less often than the molecular strategies themselves. The first is delivery to the correct cell population, since osteoblasts and skeletal stem cells are distributed through a mineralised tissue with limited vascular access, and vector tropism for these populations is imperfect. The second is durability in a growing skeleton, where dilution of episomal genomes during rapid expansion of the osteoprogenitor pool is expected to reduce expression over precisely the period when benefit is most needed. The third is the variant-specific nature of allele-directed reagents, which conflicts with a development and regulatory model built around therapies applicable to defined populations; a therapy effective for one family's variant may require substantial independent evaluation for another. The fourth is the evidentiary standard that any such therapy will face, since a durable intervention in children requires long-term safety follow-up and a fracture-based efficacy endpoint, and the surrogate endpoints on which pharmacological development has relied are unlikely to be sufficient for an irreversible intervention.

An additional consideration concerns timing. If the therapeutic window is developmental, the interventions most likely to alter the natural history are those delivered in infancy or before birth, which raises the ethical and methodological difficulties associated with intervening in a condition whose severity cannot be predicted precisely at that stage. Prenatal cell therapy trials have confronted these questions explicitly, and the same considerations will apply with greater force to irreversible genetic modification (Sagar *et al*., 2024).

Table 5 sets out the principal cell and gene-directed modalities, the highest level of evidence attained for each, the model or population studied and the specific barrier that separates each from clinical application. Read together with Table 4, it shows that the field's therapeutic ambition is currently inverted relative to its evidence: the approaches with the strongest mechanistic rationale have the weakest clinical evidence, while those with the most clinical evidence address the mechanism least directly.

**Table 5. Cell and gene-directed modalities in osteogenesis imperfecta, with highest evidence attained, study system and principal translational barrier**

| Modality | Highest level of evidence identified | Study system or population | Principal reported finding | Principal translational barrier | Supporting references |
|---|---|---|---|---|---|
| Mesenchymal stromal cell transplantation, postnatal | Case reports and case series synthesised in a systematic review | Twelve patients across nine clinical reports | Increased total body bone mineral content in five reports; reduced fracture rate in four | No controlled comparison; engraftment low and often transient | Götherström *et al*. (2021); Khorasani *et al*. (2025) |
| Mesenchymal stromal cell transplantation, prenatal and postnatal | Exploratory open-label phase I/II trial protocol | Fetuses and infants with severe disease | Trial designed for safety and tolerability | Uncontrolled design; attribution of benefit between prenatal and postnatal components | Sagar *et al*. (2024) |
| Repeated intravenous and intraosseous stromal cell dosing | Trial protocol | Fifteen children aged one to five years with severe disease | Primary aim is safety and tolerability | Small sample; efficacy endpoints are secondary and uncontrolled | Madhuri *et al*. (2025) |
| Ex vivo correction in patient-derived stem cells | Controlled laboratory experiment | Patient-derived induced pluripotent stem cells with *COL1A1* variants | Restoration of osteogenic capacity after correction | No demonstration of skeletal engraftment or in vivo benefit | Cao *et al*. (2023); Jung *et al*. (2021) |
| Promoter base editing to modulate expression | Controlled laboratory experiment | Human fibroblasts | Modulation of collagen I expression without double-strand breaks | Quantitative modulation does not address dominant-negative chains | Daliri *et al*. (2025) |

| Modality | Highest level of evidence identified | Study system or population | Principal reported finding | Principal translational barrier | Supporting references |
|---|---|---|---|---|---|
| RNA-directed rescue of a translational defect | Mechanistic study with proposed rescue strategy | Cells carrying a *SEC24D* untranslated region variant | Identification of translational dysfunction amenable to RNA-level correction | Applicable only to specific variant classes; no in vivo data | Essawi *et al*. (2026) |
| Systemic adeno-associated virus delivery of editing machinery | In vivo animal experiment | Mouse model of severe dominant disease | Reduced matrix turnover and improved osteocyte network after dual vector delivery | Editing efficiency in skeletal progenitors; durability during growth | Yang *et al*. (2024) |
| Adeno-associated virus gene transfer in recessive disease | In vivo animal experiment | Rat model of recessive disease with intraosseous delivery | Recovery of bone density and mechanical strength | Local delivery route; short follow-up; recessive mechanism only | Li *et al*. (2026) |
| Allele-specific silencing and editing platforms generally | Narrative and technical review | Not applicable | Technical feasibility described across collagen disorders | Variant-specific reagents conflict with population-based development models | Kocsy *et al*. (2025); Schindeler *et al*. (2022) |

*Note. No completed clinical trial of gene editing or gene transfer in this disorder was identified at the search date. Trial protocols are included because they define the design and endpoints of ongoing work; they do not constitute evidence of effect*

## 8. Future Directions

The priorities set out below follow from the specific weaknesses identified in the preceding synthesis rather than from a general call for further research, and they are ordered from those that are immediately actionable to those that depend on prior methodological work.

The most urgent priority is formal qualification of surrogate endpoints against fracture outcomes. The recurring pattern across pharmacological classes is that densitometric and microstructural measures respond while fracture effects remain unproven, and the field currently lacks any means of determining which surrogate changes predict clinical benefit. The appropriate approach is meta-analysis of individual participant data across completed randomised trials, testing whether treatment effects on candidate surrogates predict treatment effects on fracture within trials, combined with prospective collection of paired surrogate and fracture data in registry cohorts. Candidate surrogates should include estimated bone strength from high-resolution peripheral imaging and material-level measures, since these are mechanistically closer to fracture than areal density (Glorieux *et al*., 2024; Misof & Fratzl-Zelman, 2024). Success would allow future trials to be shorter and smaller; continued failure to attempt qualification will perpetuate the present situation in which each new agent requires a fracture-endpoint trial that the rare disease population can barely support.

The second priority is trial design adapted to a rare, heterogeneous and slowly progressing disorder. Conventional parallel-group designs powered for fracture require samples that exceed the capacity of most national populations, which is the practical reason that so many trials have used densitometric endpoints. Registry-embedded randomisation, adaptive and platform designs sharing a common control group across candidate agents, and Bayesian analysis incorporating prior information offer routes to acceptable power at feasible sample size. The trial of sequential anabolic and antiresorptive treatment demonstrates that fracture-endpoint designs are achievable in adults and should serve as a template rather than an exception (Hald *et al*., 2023). Standard outcome sets should be adopted as the default measurement framework so that trials become comparable and eligible for pooled analysis (Blokland *et al*., 2024; Goderie *et al*., 2023; Nijhuis *et al*., 2021).

The third priority is resolution of penetrance and genotype-phenotype relationships in unselected populations. Estimates derived from clinic cohorts are inflated by ascertainment, and the reduced penetrance observed in a large adult biobank has direct implications for genetic counselling and for proposals to include this disorder in genomic newborn screening (Pagnamenta *et al*., 2026). What is required is analysis of additional population-scale cohorts with skeletal phenotyping better than diagnostic coding and heel densitometry, ideally including targeted clinical assessment of variant carriers identified through sequencing. Until such work is completed, screening proposals risk identifying large numbers of individuals whose lifetime risk cannot be quantified.

The fourth priority is mechanism-stratified therapeutic evaluation. Quantitative and structural collagen defects differ in the biology that a drug must address, and recessive mechanisms differ from both, yet trials continue to enrol populations defined by clinical type or by the presence of any collagen gene variant. Prospective stratification by mechanism, with pre-specified subgroup analysis and adequate representation of each stratum, would allow genuinely different responses to be detected rather than averaged away. The observation that responses to pathway-directed therapy differed by disease type supports this approach and indicates that it is testable (Song *et al*., 2022).

The fifth priority concerns the specific obstacles to gene-directed therapy. Three problems require solution before clinical translation is realistic: efficient and selective delivery to osteoblasts and skeletal progenitors; durability of correction during skeletal growth, which for episomal vectors implies either integration, repeated administration or targeting of self-renewing progenitors; and a development pathway for variant-specific reagents that does not require full independent evaluation of each. Comparative studies delivering the same editing strategy to both dominant-negative and loss-of-function models within a single experimental system would clarify how far recessive successes generalise (Li *et al*., 2026; Yang *et al*., 2024). Natural history data of sufficient

granularity are also required to serve as external comparators for single-arm studies in ultra-rare genotypes, since randomisation will not be feasible for variants confined to individual families (Gatti *et al*., 2026).

The sixth priority is longer-term and concerns geographical and demographic representation. Most cohort, trial and registry evidence originates from a small number of high-income countries in Europe, North America and East Asia, and genotype spectra reported from other settings differ in the proportion of recessive disease and in variant distribution (Selina *et al*., 2025). Building cohorts and registries in under-represented regions would improve the external validity of prevalence estimates, clarify whether treatment responses differ across populations, and address an equity gap that current evidence cannot even quantify. Given the growth of high-consanguinity recessive disease in some populations, this priority also has direct relevance to gene-directed therapy, where recessive mechanisms are the more tractable target.

## 9. Conclusions

Osteogenesis imperfecta has been transformed over four decades from a clinically defined syndrome into a molecularly stratified group of disorders, and the review objectives were to assess whether classification, mechanism, outcome measurement and therapeutic evidence have advanced coherently alongside that transformation. They have not.

Classification is the clearest instance of divergence. Several frameworks now coexist, each internally consistent and each defensible on its own terms, yet none has been evaluated prospectively against prognosis, surveillance, therapeutic selection or trial stratification. The most defensible conclusion is that current classification supports communication reliably and clinical decision-making only weakly, and that claims to the contrary rest on plausibility rather than evidence. The numbered system in widest use encodes discovery order rather than mechanism or severity, and its continued use in trial eligibility criteria introduces heterogeneity that complicates cross-study comparison.

Mechanistic understanding is the strongest element of the field. The distinction between quantitative deficiency and dominant-negative structural defect is robust, well replicated and therapeutically meaningful, and the extension of causation to collagen modification, folding, secretion and osteoblast signalling is supported by convergent evidence from multiple research groups. Confidence should nevertheless be calibrated to the underlying evidence for each mechanism, which ranges from thoroughly characterised pathways to assignments resting on a single family and one functional assay. The claim that heterogeneous upstream defects converge on a shared final pathway is plausible and partially supported, not demonstrated.

The most consequential finding of this synthesis concerns the relationship between measurement and mechanism. Bone in this disorder is abnormally highly mineralised while remaining mechanically inferior, which means that mineral density is not a straightforward proxy for fracture resistance. This single observation explains a pattern that recurs across every pharmacological class examined: densitometric endpoints respond consistently while fracture endpoints do not. Antiresorptive treatment reliably increases bone mineral density, and the evidence that it reduces fractures is inconsistent and of limited certainty. Anabolic and pathway-directed agents have advanced through development on the strength of surrogate responses whose predictive validity has never been established. The lesson is methodological rather than pharmacological, and it applies to every agent that follows.

Cell and gene-directed therapy occupies a position that is frequently misrepresented. The mechanistic rationale is sound, delivery of editing machinery to the skeleton has been achieved in animals, and correction has been demonstrated in patient-derived cells. The clinical evidence, by contrast, consists of a small number of uncontrolled cell transplantation reports involving a very small total number of patients, together with exploratory trial protocols. No completed clinical trial of gene editing or gene transfer in this disorder was identified. The most defensible statement is that these approaches are biologically promising and clinically unproven, and that the dominant-negative mechanism responsible for most disease remains the hardest target rather than the one closest to solution.

The broader significance of this synthesis extends beyond a single rare disorder. Osteogenesis imperfecta illustrates with unusual clarity what happens when molecular understanding advances faster than the measurement framework required to test its therapeutic implications. Mechanism has been elucidated in detail; classification has multiplied without validation; endpoints have remained convenient rather than valid; and therapeutic development has proceeded on surrogates that mechanism itself predicts to be unreliable. Progress in the next decade will depend less on identifying additional causative genes than on qualifying endpoints, designing trials that a rare disease population can support, and holding claims about gene-directed therapy to the evidentiary standard that irreversible intervention in children demands.

## 10. Limitations

This article is a critical narrative review, and the limitations intrinsic to that design apply in full. Literature selection was purposive and iterative rather than exhaustive, appraisal was conducted without formal instruments, and synthesis depended on interpretive judgement. Selection and interpretive bias cannot be excluded despite a stated search strategy, explicit eligibility rules and transparent appraisal criteria, and a different reviewer working from the same sources could reasonably weight the evidence differently, particularly in the assessment of competing classification frameworks and in the reading of preclinical gene therapy work.

Access was limited to freely available scholarly sources, so evidence indexed only in subscription databases may have been missed, and the resulting selection may under-represent literature that is poorly covered by open indexes. Only publications in English were included, which is likely to have reduced representation of research published in other languages, particularly from regions whose

scholarly output is not comprehensively indexed internationally. Grey literature, unpublished data and conference material were excluded, and the resulting evidence base is therefore susceptible to publication bias, which is likely to have inflated the apparent effect of interventions and to have suppressed null and negative preclinical findings.

Several features of the underlying literature constrain what any synthesis can conclude. Study designs are highly heterogeneous, spanning randomised trials, observational cohorts, registry linkage, case series, animal experiments and cell models, and no common effect measure links them, so quantitative synthesis was neither attempted nor appropriate. Outcome definitions differ across studies to a degree that limits comparability even within a single therapeutic class. Clinical cohorts are drawn predominantly from specialist referral centres, which systematically over-represents severe disease and inflates estimates of penetrance, complication frequency and healthcare utilisation. Geographical representation is uneven, with most evidence originating from a small number of high-income countries, so conclusions about genotype distribution, treatment response and disease burden may not generalise to other settings.

Evidence on long-term outcomes is scarce across every domain examined. Trials of pharmacological agents have generally been short relative to the duration of clinical exposure, no controlled evidence addresses the consequences of prolonged treatment through skeletal growth, and follow-up in preclinical gene-directed studies is brief relative to the human growth period that such therapies would have to span. Much of the evidence on extraskeletal manifestations is observational or based on expert consensus rather than on controlled comparison, and several care recommendations therefore rest on opinion.

The field is also developing rapidly, and evidence relevant to several conclusions was in progress at the time of writing. In particular, the definitive evaluation of sclerostin inhibition had reached primary completion without a peer-reviewed report of its outcome being identifiable in the accessible literature, so conclusions about that therapeutic class are provisional and may require substantial revision once the results are published in full.

## Consent

It is not applicable.

## Ethical Approval

It is not applicable.

## Declaration of AI Use

This manuscript was prepared through the intellectual contributions of the author(s) to the study design, data analysis, interpretation of results, and scholarly content. Grammarly and ChatGPT were used solely to assist with grammatical refinement and reference formatting during manuscript revision. The author(s) accept full responsibility for the accuracy, integrity, and final content of the manuscript.

## Competing Interests

Authors have declared that they have no known competing financial interests or non-financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

Baethge, C., Goldbeck-Wood, S., & Mertens, S. (2019). SANRA—a scale for the quality assessment of narrative review articles. *Research Integrity and Peer Review, 4*, 5. https://doi.org/10.1186/s41073-019-0064-8

Batkovskyte, D., Swolin-Eide, D., Hammarsjö, A., Sæther, K. B., Thunström, S., Lundin, J., Eisfeldt, J., Lindstrand, A., Nordgren, A., Åström, E., & Grigelioniene, G. (2025). Structural variants in *COL1A1* and *COL1A2* in osteogenesis imperfecta. *American Journal of Medical Genetics Part A, 197*(3), e63935. https://doi.org/10.1002/ajmg.a.63935

Besio, R., Contento, B. M., Garibaldi, N., Filibian, M., Sonntag, S., Shmerling, D., Tonelli, F., Biggiogera, M., Brini, M., Salmaso, A., Jovanovic, M., Marini, J. C., Rossi, A., & Forlino, A. (2023). CaMKII inhibition due to TRIC-B loss-of-function dysregulates SMAD signaling in osteogenesis imperfecta. *Matrix Biology, 120*, 43–59. https://doi.org/10.1016/j.matbio.2023.05.002

Blandin, C., Collet, C., Ostertag, A., Funck-Brentano, T., & Cohen-Solal, M. (2025). Genetics and bone mineral density predict the fractures in adults with osteogenesis imperfecta: A prospective study. *The Journal of Clinical Endocrinology & Metabolism, 110*(8), 2307–2313. https://doi.org/10.1210/clinem/dgae791

Blokland, L., Arponen, H., Ahmad, A., Colijn, S., Gjørup, H., John, R., Li, M., Mekking, D., Parekh, S., Retrouvey, J. M., Stutz Steiger, T., Zhou, L., & Andersson, K. (2024). A standard set of outcome measures for the comprehensive assessment of oral health and occlusion in individuals with osteogenesis imperfecta. *Orphanet Journal of Rare Diseases, 19*, 294. https://doi.org/10.1186/s13023-024-03308-5

Cabral, W. A., Chang, W., Barnes, A. M., Weis, M., Scott, M. A., Leikin, S., Makareeva, E., Kuznetsova, N. V., Rosenbaum, K. N., Tifft, C. J., Bulas, D. I., Kozma, C., Smith, P. A., Eyre, D. R., & Marini, J. C. (2007). Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. *Nature Genetics, 39*(3), 359–365. https://doi.org/10.1038/ng1968

Cao, Y., Li, L., Ren, X., Mao, B., Yang, Y., Mi, H., Guan, Y., Li, S., Zhou, S., Guan, X., Yang, T., & Zhao, X. (2023). CRISPR/Cas9 correction of a dominant cis-double-variant in *COL1A1* isolated from a patient with osteogenesis imperfecta increases the osteogenic capacity of induced pluripotent stem cells. *Journal of Bone and Mineral Research, 38*(5), 719–732. https://doi.org/10.1002/jbmr.4783

Chang, W., Barnes, A. M., Cabral, W. A., Bodurtha, J. N., & Marini, J. C. (2010). Prolyl 3-hydroxylase 1 and CRTAP are mutually stabilizing in the endoplasmic reticulum collagen prolyl 3-hydroxylation complex. *Human Molecular Genetics, 19*(2), 223–234. https://doi.org/10.1093/hmg/ddp481

Chaugule, S., Constantinou, C. K., John, A. A., Micha, D., Eekhoff, M., Gravallese, E., Gao, G., & Shim, J.-H. (2025). Comprehensive review of osteogenesis imperfecta: Current treatments and future innovations. *Human Gene Therapy, 36*(5–6), 597–617. https://doi.org/10.1089/hum.2024.191

Chetty, M., Roomaney, I. A., & Beighton, P. (2021). The evolution of the nosology of osteogenesis imperfecta. *Clinical Genetics, 99*(1), 42–52. https://doi.org/10.1111/cge.13846

Dalgleish, R., Micha, D., Superti-Furga, A., van Dijk, F. S., & Sillence, D. O. (2024). Letter to the editor: Re: Pathogenic mechanisms of osteogenesis imperfecta, evidence for classification. *Orphanet Journal of Rare Diseases, 19*, 272. https://doi.org/10.1186/s13023-024-03269-9

Daliri, K., Hescheler, J., Newby, G. A., Clement, K., Liu, D. R., & Pfannkuche, K. (2025). Modulating collagen I expression in fibroblasts by CRISPR-Cas9 base editing of the collagen 1A1 promoter. *International Journal of Molecular Sciences, 26*(7), 3041. https://doi.org/10.3390/ijms26073041

Dwan, K., Phillipi, C. A., Steiner, R. D., & Basel, D. (2016). Bisphosphonate therapy for osteogenesis imperfecta. *Cochrane Database of Systematic Reviews, 2016*(10), CD005088. https://doi.org/10.1002/14651858.CD005088.pub4

Essawi, O., Jarayseh, T., Tapaneeyaphan, P., Bouckaert, M., Naessens, S., Malfait, F., Coppieters, F., & Coucke, P. J. (2026). A rare 5′UTR variant in *SEC24D* reveals translational dysfunction in osteogenesis imperfecta: A roadmap for RNA therapeutic rescue. *Scientific Reports, 16*, 687. https://doi.org/10.1038/s41598-025-29937-9

Folkestad, L., Prakash, S. K., Nagamani, S. C. S., Andersen, N. H., Carter, E., Hald, J. D., Johnson, R. J., Langdahl, B., Perfetto, E. M., Raggio, C., Ralston, S. H., Sandhaus, R. A., Semler, O., Tosi, L., & Orwoll, E. (2025). Cardiovascular disease in adults with osteogenesis imperfecta: Clinical characteristics, care recommendations, and research priorities identified using a modified Delphi technique. *Journal of Bone and Mineral Research, 40*(2), 211–221. https://doi.org/10.1093/jbmr/zjae197

Forlino, A., & Marini, J. C. (2016). Osteogenesis imperfecta. *The Lancet, 387*(10028), 1657–1671. https://doi.org/10.1016/S0140-6736(15)00728-X

Garibaldi, N., Besio, R., Dalgleish, R., Villani, S., Barnes, A. M., Marini, J. C., & Forlino, A. (2022). Dissecting the phenotypic variability of osteogenesis imperfecta. *Disease Models & Mechanisms, 15*(5), dmm049398. https://doi.org/10.1242/dmm.049398

Gatti, D., Prince, S., Sazova, O., Whitcher, C., Randet, A., Carter, M., Rapoport, M., & Semler, O. (2026). The natural history of osteogenesis imperfecta: A systematic review. *Bone Reports, 29*, 101927. https://doi.org/10.1016/j.bonr.2026.101927

Ghosh, D. K., Udupa, P., Shrikondawar, A. N., Bhavani, G. S., Shah, H., Ranjan, A., & Girisha, K. M. (2023). Mutant MESD links cellular stress to type I collagen aggregation in osteogenesis imperfecta type XX. *Matrix Biology, 115*, 81–106. https://doi.org/10.1016/j.matbio.2022.12.001

Glorieux, F. H., Bishop, N. J., Plotkin, H., Chabot, G., Lanoue, G., & Travers, R. (1998). Cyclic administration of pamidronate in children with severe osteogenesis imperfecta. *New England Journal of Medicine, 339*(14), 947–952. https://doi.org/10.1056/NEJM199810013391402

Glorieux, F. H., Langdahl, B., Chapurlat, R., De Beur, S. J., Sutton, V. R., Poole, K. E. S., Dahir, K. M., Orwoll, E. S., Willie, B. M., Mikolajewicz, N., Zimmermann, E., Hosseinitabatabaei, S., Ominsky, M. S., Saville, C., Clancy, J., MacKinnon, A., Mistry, A., & Javaid, M. K. (2024). Setrusumab for the treatment of osteogenesis imperfecta: 12-month results from the phase 2b asteroid study. *Journal of Bone and Mineral Research, 39*(9), 1215–1228. https://doi.org/10.1093/jbmr/zjae112

Gochuico, B. R., Hossain, M., Talvacchio, S. K., Zuo, M. X. G., Barton, M., Dang Do, A. N., & Marini, J. C. (2023). Pulmonary function and structure abnormalities in children and young adults with osteogenesis imperfecta point to intrinsic and extrinsic lung abnormalities. *Journal of Medical Genetics, 60*(11), 1067–1075. https://doi.org/10.1136/jmg-2022-109009

Goderie, T., Hendricks, S., Cocchi, C., Maroger, I. D., Mekking, D., Mosnier, I., Musacchio, A., Vernick, D., & Smits, C. (2023). The international standard set of outcome measures for the assessment of hearing in people with osteogenesis imperfecta. *Otology & Neurotology, 44*(7), e449–e455. https://doi.org/10.1097/MAO.0000000000003921

Götherström, C., David, A. L., Walther-Jallow, L., Åström, E., & Westgren, M. (2021). Mesenchymal stem cell therapy for osteogenesis imperfecta. *Clinical Obstetrics & Gynecology, 64*(4), 898–903. https://doi.org/10.1097/GRF.0000000000000656

Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation, 48*(6), e12931. https://doi.org/10.1111/eci.12931

Hald, J. D., Keerie, C., Weir, C. J., Javaid, M. K., Lam, W., Osborne, P., Walsh, J., Langdahl, B. L., & Ralston, S. H. (2023). Protocol of a randomised trial of teriparatide followed by zoledronic acid to reduce fracture risk in adults with osteogenesis imperfecta. *BMJ Open, 13*(11), e078164. https://doi.org/10.1136/bmjopen-2023-078164

Hasegawa, K. (2025). Osteogenesis imperfecta: Pathogenesis, classification, and treatment. *Clinical Pediatric Endocrinology, 34*(3), 152–161. https://doi.org/10.1297/cpe.2025-0009

Haumann, S. K., Sørensen, J. R., Schmidt, J. H., & Folkestad, L. (2024). The PATCH study: Prevalence of hearing loss during ageing and treatment choices in osteogenesis imperfecta: A Danish nationwide register-based cohort study. *Calcified Tissue International, 115*(3), 260–268. https://doi.org/10.1007/s00223-024-01253-w

Jovanovic, M., & Marini, J. C. (2024). Update on the genetics of osteogenesis imperfecta. *Calcified Tissue International, 115*(6), 891–914. https://doi.org/10.1007/s00223-024-01266-5

Jovanovic, M., Guterman-Ram, G., & Marini, J. C. (2022). Osteogenesis imperfecta: Mechanisms and signaling pathways connecting classical and rare OI types. *Endocrine Reviews, 43*(1), 61–90. https://doi.org/10.1210/endrev/bnab017

Jung, H., Rim, Y. A., Park, N., Nam, Y., & Ju, J. H. (2021). Restoration of osteogenesis by CRISPR/Cas9 genome editing of the mutated *COL1A1* gene in osteogenesis imperfecta. *Journal of Clinical Medicine, 10*(14), 3141. https://doi.org/10.3390/jcm10143141

Khorasani, M. M., Ebrahimzadeh, M. H., Dehghani, M., Sharafoddin, M., Moradi, A., & Jirofti, N. (2025). Mesenchymal stem cell transplantation for osteogenesis imperfecta patients: A systematic review. *Annals of the New York Academy of Sciences, 1554*(1), 27–44. https://doi.org/10.1111/nyas.70066

Kocsy, K., Wilkinson, H., Felix-Ilemhenbhio, F., Bax, B., Van Agtmael, T., Azzouz, M., & Majid, A. (2025). Gene editing for collagen disorders: Current advances and future perspectives. *Gene Therapy, 32*(6), 676–689. https://doi.org/10.1038/s41434-025-00560-7

Li, S., Chen, X., Cao, Y., Han, M., Guan, F., Ren, X., Mi, H., Yang, T., Li, M., & Zhao, X. (2026). A new *Wnt1* mutant rat model of osteogenesis imperfecta and its application in AAV9-mediated gene therapy. *Human Mutation, 2026*, 7351808. https://doi.org/10.1155/humu/7351808

Liu, J., Lin, X., Sun, L., Zhang, Q., Jiang, Y., Wang, O., Xing, X., Xia, W., & Li, M. (2024). Safety and efficacy of denosumab in children with osteogenesis imperfecta—The first prospective comparative study. *The Journal of Clinical Endocrinology & Metabolism, 109*(7), 1827–1836. https://doi.org/10.1210/clinem/dgad732

Liu, W., Nicol, L., & Orwoll, E. (2024). Current and developing pharmacologic agents for improving skeletal health in adults with osteogenesis imperfecta. *Calcified Tissue International, 115*(6), 805–811. https://doi.org/10.1007/s00223-024-01188-2

Madhuri, V., Ramesh, S., Goos, A., Paul, T. V., Nidugala Kesava, S., Mathews, V., Walther-Jallow, L., & Götherström, C. (2025). Evaluation of safety and efficacy of multiple intravenous and intraosseous doses of foetal liver-derived mesenchymal stem cells in children with severe osteogenesis imperfecta: The BOOST2B clinical trial protocol. *Bone & Joint Open, 6*(3), 361–372. https://doi.org/10.1302/2633-1462.63.BJO-2024-0115.R1

Marini, J. C., Forlino, A., Bächinger, H. P., Bishop, N. J., Byers, P. H., De Paepe, A., Fassier, F., Fratzl-Zelman, N., Kozloff, K. M., Krakow, D., Montpetit, K., & Semler, O. (2017). Osteogenesis imperfecta. *Nature Reviews Disease Primers, 3*, 17052. https://doi.org/10.1038/nrdp.2017.52

Marom, R., Song, I. W., Busse, E. C., Washington, M. E., Berrier, A. S., Rossi, V. C., Ortinau, L., Jeong, Y., Jiang, M. M., Dawson, B. C., Adeyeye, M., Leynes, C., Lietman, C. D., Stroup, B. M., Batkovskyte, D., Jain, M., Chen, Y., Cela, R., Castellon, A., … Lee, B. H. (2024). The *IFITM5* mutation in osteogenesis imperfecta type V is associated with an ERK/SOX9-dependent osteoprogenitor differentiation defect. *Journal of Clinical Investigation, 134*(15), e170369. https://doi.org/10.1172/JCI170369

Marulanda, J., Tauer, J. T., Boraschi-Diaz, I., Bardai, G., & Rauch, F. (2023). Effect of sclerostin inactivation in a mouse model of severe dominant osteogenesis imperfecta. *Scientific Reports, 13*, 5010. https://doi.org/10.1038/s41598-023-32221-3

Misof, B. M., & Fratzl-Zelman, N. (2024). Bone quality and mineralization and effects of treatment in osteogenesis imperfecta. *Calcified Tissue International, 115*(6), 777–804. https://doi.org/10.1007/s00223-024-01263-8

Misof, B. M., Roschger, P., Mähr, M., Fratzl-Zelman, N., Glorieux, F. H., Hartmann, M. A., Rauch, F., & Blouin, S. (2023). Accelerated mineralization kinetics in children with osteogenesis imperfecta type 1. *Bone, 166*, 116580. https://doi.org/10.1016/j.bone.2022.116580

Munns, C. F., Rauch, F., Zeitlin, L., Fassier, F., & Glorieux, F. H. (2004). Delayed osteotomy but not fracture healing in pediatric osteogenesis imperfecta patients receiving pamidronate. *Journal of Bone and Mineral Research, 19*(11), 1779–1786. https://doi.org/10.1359/JBMR.040814

Nijhuis, W., Franken, A., Ayers, K., Damas, C., Folkestad, L., Forlino, A., Fraschini, P., Hill, C., Janus, G., Kruse, R., Lande Wekre, L., Michiels, L., Montpetit, K., Panzeri, L., Porquet-Bordes, V., Rauch, F., Sakkers, R., Salles, J. P., Semler, O., … Verhoef, M. (2021). A standard set of outcome measures for the comprehensive assessment of osteogenesis imperfecta. *Orphanet Journal of Rare Diseases, 16*, 140. https://doi.org/10.1186/s13023-021-01682-y

Ning, H., Liang, C., Mei, H., Yuan, D., Wei, X., Huang, X., Tan, D., & Tan, J. (2025). A novel homozygous synonymous variant in *CCDC134* as a cause of osteogenesis imperfecta type XXII. *Clinical Genetics, 107*(4), 446–452. https://doi.org/10.1111/cge.14664

Pagnamenta, A. T., Fasham, J., Beaumont, R. N., Baker, D., Keigwin, S., Hall, T., Baple, E. L., Balasubramanian, M., Jackson, L., & Wright, C. F. (2026). Reduced penetrance of *COL1A1/2* pathogenic variants linked with osteogenesis imperfecta: Analysis of a large population cohort. *European Journal of Human Genetics*. Advance online publication. https://doi.org/10.1038/s41431-026-02118-6

Panzaru, M. C., Florea, A., Caba, L., & Gorduza, E. V. (2023). Classification of osteogenesis imperfecta: Importance for prophylaxis and genetic counseling. *World Journal of Clinical Cases, 11*(12), 2604–2620. https://doi.org/10.12998/wjcc.v11.i12.2604

Philippoteaux, C., Lefevre, C., Saint-Dizier, C., Paccou, J., & Chazard, E. (2025). Mortality and fracture risk in children with osteogenesis imperfecta: Results from the French nationwide hospital discharge database. *Bone, 200*, 117607. https://doi.org/10.1016/j.bone.2025.117607

Prado, H. V., Soares, E. C. B., Carneiro, N. C. R., Vilar, I. C. O., Abreu, L. G., & Borges-Oliveira, A. C. (2023). Dental anomalies in individuals with osteogenesis imperfecta: A systematic review and meta-analysis of prevalence and comparative studies. *Journal of Applied Oral Science, 31*, e20230040. https://doi.org/10.1590/1678-7757-2023-0040

Rauch, F., & Glorieux, F. H. (2004). Osteogenesis imperfecta. *The Lancet, 363*(9418), 1377–1385. https://doi.org/10.1016/S0140-6736(04)16051-0

Rummler, M., Schemenz, V., McCluskey, S., Davydok, A., Rauch, F., Glorieux, F. H., Harrington, M. J., Wagermaier, W., Willie, B. M., & Zimmermann, E. A. (2024). Bone matrix properties in adults with osteogenesis imperfecta are not adversely affected by setrusumab—a sclerostin neutralizing antibody. *Journal of Bone and Mineral Research, 39*(9), 1229–1239. https://doi.org/10.1093/jbmr/zjae108

Sagar, R. L., Åström, E., Chitty, L. S., Crowe, B., David, A. L., DeVile, C., Forsmark, A., Franzen, V., Hermeren, G., Hill, M., Johansson, M., Lindemans, C., Lindgren, P., Nijhuis, W., Oepkes, D., Rehberg, M., Sahlin, N. E., Sakkers, R., Semler, O.,

… Götherström, C. (2024). An exploratory open-label multicentre phase I/II trial evaluating the safety and efficacy of postnatal or prenatal and postnatal administration of allogeneic expanded fetal mesenchymal stem cells for the treatment of severe osteogenesis imperfecta in infants and fetuses: The BOOSTB4 trial protocol. *BMJ Open, 14*(6), e079767. https://doi.org/10.1136/bmjopen-2023-079767

Sait, H., Adarsha, N., Moirangthem, A., Saxena, D., Sharma, L., Dabadgao, P., Gupta, A., & Phadke, S. R. (2025). Molecular and clinical landscape of osteogenesis imperfecta: Unraveling autosomal recessive forms, therapeutic outcomes, and bone mineral density in carriers. *Clinical Genetics, 108*(6), 684–695. https://doi.org/10.1111/cge.70003

Schindeler, A., Lee, L. R., O'Donohue, A. K., Ginn, S. L., & Munns, C. F. (2022). Curative cell and gene therapy for osteogenesis imperfecta. *Journal of Bone and Mineral Research, 37*(5), 826–836. https://doi.org/10.1002/jbmr.4549

Sclocco, A., Yang, W., Ventura, L., Dubois, L., Eekhoff, E. M. W., Micha, D., & Zhytnik, L. (2026). Dose matters: Haploinsufficiency in osteogenesis imperfecta. *Nature Reviews Endocrinology*. Advance online publication. https://doi.org/10.1038/s41574-026-01282-5

Selina, A., Kandagaddala, M., Thomas, N., Paul, T., Chapla, A., Danda, S., & Madhuri, V. (2025). *COL1A1* and *COL1A2* gene variants causing osteogenesis imperfecta in a major referral center of India. *American Journal of Medical Genetics Part A, 197*(7), e64023. https://doi.org/10.1002/ajmg.a.64023

Sillence, D. O. (2024). A dyadic nosology for osteogenesis imperfecta and bone fragility syndromes 2024. *Calcified Tissue International, 115*(6), 873–890. https://doi.org/10.1007/s00223-024-01248-7

Sillence, D. O., Senn, A., & Danks, D. M. (1979). Genetic heterogeneity in osteogenesis imperfecta. *Journal of Medical Genetics, 16*(2), 101–116. https://doi.org/10.1136/jmg.16.2.101

Song, I. W., Nagamani, S. C., Nguyen, D., Grafe, I., Sutton, V. R., Gannon, F. H., Munivez, E., Jiang, M. M., Tran, A., Wallace, M., Esposito, P., Musaad, S., Strudthoff, E., McGuire, S., Thornton, M., Shenava, V., Rosenfeld, S., Huang, S., Shypailo, R., … Lee, B. (2022). Targeting TGF-β for treatment of osteogenesis imperfecta. *Journal of Clinical Investigation, 132*(7), e152571. https://doi.org/10.1172/JCI152571

Storoni, S., Treurniet, S., Maugeri, A., Pals, G., van den Aardweg, J. G., van der Pas, S. L., Elting, M. W., Kloen, P., Micha, D., & Eekhoff, E. M. W. (2022). Prevalence and hospital admissions in patients with osteogenesis imperfecta in the Netherlands: A nationwide registry study. *Frontiers in Endocrinology, 13*, 869604. https://doi.org/10.3389/fendo.2022.869604

Storoni, S., Verdonk, S. J. E., Zhytnik, L., Pals, G., Treurniet, S., Elting, M. W., Sakkers, R. J. B., van den Aardweg, J. G., Eekhoff, E. M. W., & Micha, D. (2023). From genetics to clinical implications: A study of 675 Dutch osteogenesis imperfecta patients. *Biomolecules, 13*(2), 281. https://doi.org/10.3390/biom13020281

Sukhera, J. (2022). Narrative reviews: Flexible, rigorous, and practical. *Journal of Graduate Medical Education, 14*(4), 414–417. https://doi.org/10.4300/JGME-D-22-00480.1

Tan, Z., Shek, H. T., Li, Z., Xia, L., He, Y., Chen, P., Wong, J. S. H., Gao, B., Chan, D., & To, M. K. T. (2025). An inducible mouse model of osteogenesis imperfecta type V reveals aberrant osteogenesis caused by *Ifitm5* c.-14C>T mutation. *Journal of Bone and Mineral Research, 40*(5), 577–590. https://doi.org/10.1093/jbmr/zjaf022

U.S. National Library of Medicine. (2026, February 25). *Setrusumab vs placebo for osteogenesis imperfecta (Orbit)* (ClinicalTrials.gov Identifier No. NCT05125809). ClinicalTrials.gov. https://clinicaltrials.gov/study/NCT05125809

Unger, S., Ferreira, C. R., Mortier, G. R., Ali, H., Bertola, D. R., Calder, A., Cohn, D. H., Cormier-Daire, V., Girisha, K. M., Hall, C., Krakow, D., Makitie, O., Mundlos, S., Nishimura, G., Robertson, S. P., Savarirayan, R., Sillence, D., Simon, M., Sutton, V. R., … Superti-Furga, A. (2023). Nosology of genetic skeletal disorders: 2023 revision. *American Journal of Medical Genetics Part A, 191*(5), 1164–1209. https://doi.org/10.1002/ajmg.a.63132

van Welzenis, T., Westerheim, I., Hart, T., Wekre, L. L., Semler, O., Rauch, F., Dewavrin, L., Dadzie, R., Prince, S., & Raggio, C. (2024). The IMPACT survey: The humanistic impact of osteogenesis imperfecta in adults. *BMC Public Health, 24*, 3318. https://doi.org/10.1186/s12889-024-20555-0

Wang, Y. Y., Su, Y. C., Lai, P. C., Chou, Y. Y., Wu, P. T., Tsai, M. C., Tai, T. W., Wu, C. H., Chang, Y. F., Tu, Y. K., Fang, C. J., Lin, C. J., Kuan, F. C., Hsu, K. L., Hong, C. K., Su, W. R., Huang, M. T., & Shih, C. A. (2025). Optimizing bone health with bisphosphonate therapies in pediatric osteogenesis imperfecta: A network meta-analysis of randomized trials. *Archives of Osteoporosis, 20*(1), 33. https://doi.org/10.1007/s11657-025-01515-6

Yang, Y. S., Sato, T., Chaugule, S., Ma, H., Xie, J., Gao, G., & Shim, J. H. (2024). AAV-based gene editing of type 1 collagen mutation to treat osteogenesis imperfecta. *Molecular Therapy Nucleic Acids, 35*(1), 102111. https://doi.org/10.1016/j.omtn.2023.102111

Yu, H., Li, C., Wu, H., Xia, W., Wang, Y., Zhao, J., & Xu, C. (2023). Pathogenic mechanisms of osteogenesis imperfecta, evidence for classification. *Orphanet Journal of Rare Diseases, 18*, 234. https://doi.org/10.1186/s13023-023-02849-5

Zervou, Z., Bruggenwirth, H. T., Demirdas, S., & Zillikens, M. C. (2025). Phenotypic and genetic characteristics of a Dutch cohort of patients with X-linked osteoporosis due to *PLS3* genetic variants. *JBMR Plus, 9*(6), ziaf046. https://doi.org/10.1093/jbmrpl/ziaf046



*Peer-review history:*
*The peer review history for this paper can be accessed here:*
*https://pr.sdiarticle5.com/review-history/163511*